# Comparing Primitive Concept Fields in Clinical Terminologies

Colin Price MPhil FRCS, Philip JB Brown MRCGP, Erich B Schulz MB BS
NHS Centre for Coding and Classification,
Loughborough, United Kingdom.

## INTRODUCTION

Standardisation of clinical vocabularies is the focus of ongoing attention and an important component of such convergence will be an agreed set of primitive concepts for populating compatible concept models.

The United Kingdom National Health Service Centre for Coding and Classification maintains a number of clinical terminology systems including all versions of the Read Codes and also the UK Surgical Procedure Classification (OPCS4).

Using object-attribute-value triples, we have semantically defined parallel sections of the Read Thesaurus (a natural language vocabulary) and OPCS4 (a formal classification) in the domain of cardiac surgical procedures. The technique is described and parameters for comparison of results and standardisation are presented.

## TECHNIQUE

Concepts from both terminologies were examined to identify intrinsic characteristics (eg site. method, approach). Using these attributes, a frame-based characterisation of each object concept was then performed to identify appropriate values (table 1):

Table 1: Semantic definition

| *Object:* **Open pulmonary valvectomy** | |
|---|---|
| *Attribute* | *Value* |
| Site | Pulmonary valve |
| Method | Excision |
| Approach | Open |

On completion of this process, the applicable values for each attribute were extracted and compared.

## PARAMETERS

For each terminology:

- $C_n$:. *The total set of object concepts.*

For each attribute:

- $T_n$: *The set of pre-coordinated object concepts in the terminology in which it is represented.*

For each concept field:

- K: *The total set of discrete primitive values.* $K = V_1 \cup V_2$.
- $V_n$: *The subset of K applied to one terminology.*
- $U_n$: *The subset of K unique to one terminology.* $U_1 = V_1 - X$.
- X: *The subset of K shared between the terminologies.* $X = V_1 \cap V_2$.

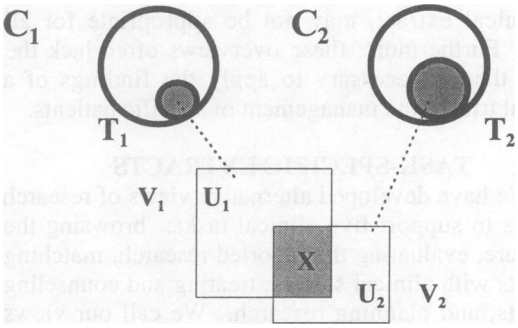


Figure 1: Illustration of parameters

## RESULTS

Table 2: Results for attribute *Approach*

| | Read (=C1) | OPCS4 (=C2) |
|---|---|---|
| C | 688 | 445 |
| T | 87 | 112 |
| K | 20 | |
| V | 14 | 11 |
| U | 9 | 6 |

## COMMENT

Representation of complex concepts involves the specification of primitive components and standardisation of these is fundamental to the development of a shared ontology. Such standardisation in the test domain (figure 1) requires:

- No ambiguity in the concepts in "X".
- Reconciliation of duplicates in $V_1$ and $V_2$.

1091—8280/97/$5.00 